DOCUMENT RESUME

ED 161 069 TM 009 987

AUTHOR Murray, Stephen L.
TITLE An Analysis of Regression Effects and the Equipercentile Growth Assumption in the Norm-Referenced Evaluation Model.
PUB DATE May 78
NOTE 30p.: Paper presented at the Annual Meeting of the Washington Educational Research Association (Seattle, WA, May 25-26, 1978)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Achievement Gains: Educational Assessment: Elementary Secondary Education: Evaluation Methods: *Mathematical Models: National Norms: *Norm Referenced Tests: Pretests: Program Evaluation *Research Design: Research Problems: Scores: *Statistical Analysis: Statistical Bias: Technical Reports
IDENTIFIERS Elementary Secondary Education Act Title I: *Regression (Statistical): *RMC Models

ABSTRACT

The norm-referenced evaluation model (RMC Model A) for Title I project evaluation, consists of procedures whereby the expected posttest standing of a treatment group under the null condition is generated from their pretest standing. It is assumed that the treatment group is not selected on the basis of their pretest scores and can be considered representative of the population represented in the test norms. Given these assumptions, the Model A rule for estimating the expected no-treatment posttest status is that a comparison group given no special treatment will maintain its percentile across time (the equipercentile growth assumption). This paper provides a theoretical analysis of the no-treatment expectation estimation procedures for Model A. An alternative definition of the no-treatment expectation is based on a structural equation model and statistical models underlying nonequivalent control group designs. Results indicate that using a measure other than the pretest score for selection reduces but does not control regression effects. If the selection measure is more highly correlated with the pretest measure than with the posttest measure, additional regression occurs between the pretest and the posttest. Biased measures of treatment effects may results. (Author/CTM)

ED181069

AN ANALYSIS OF REGRESSION EFFECTS AND THE EQUIPERCENTILE GROWTH ASSUMPTION IN THE NORM-REFERENCED EVALUATION MODEL

Stephen L. Murray
Northwest Regional Educational Laboratory
Title I Evaluation Technical Assistance Center

May 1978

Paper presented at the Annual Meeting of the Washington Educational Research Association, Seattle, Washington, May 25-26, 1978

TM009 987

In 1974 Congress amended the law governing ESEA Title I project evaluation to require comparable data across projects. These data are to include objective measures of student achievement. As a result of the 1974 amendment, the United States Office of Education (USOE) initiated development of the USOE Title I Evaluation and Reporting System. The System, developed by the RMC Research Corporation, applies to Title I projects in basic skill areas of mathematics, reading and language arts in grades 2-12. Implementation of the USOE System would standardize Title I evaluation procedures to provide comparable evidence concerning Title I project effects on student achievement.

At the heart of the System are three alternative evaluation models. A local project may choose from these models or develop an acceptable alternative to generate Title I effects data needed at the Federal level. The three models are: the <u>Control Group Model</u> (Model B), the <u>Special Regression Model</u> (Model C) and the <u>Norm-Referenced Model</u> (Model A).

In describing each of the models RMC indicates that, if appropriately implemented, they will produce unbiased estimates of a Title I project <u>treatment effect</u> (Tallmadge and Wood, 1976). Each model provides a rationale for quantifying the expected mean achievement score of the Title I project students at the termination of a project period under the assumption that the project has had <u>no effect at all</u>. This score, called the <u>no-treatment expectation</u>, is subtracted from the observed mean posttreatment achievement score of Title I students, resulting in the Title I project treatment effect.

The most popular model among practitioners is the Norm-Referenced Model. Its popularity is probably due to ease of implementation and the fact that it has little significant effect on the student selection process. The Norm-Referenced Model requires neither random assignment of students into the Title I program nor the use of a strict cutoff score on a selection measure.

Selecting students is an important contact point between the operation of Title I projects and the implementation of Title I evaluation. Students who are in greatest need of services are the highest priority students from the point of view of program operation. On the other hand, the process of selecting students into a treatment group must be clearly defined in order to avoid bias in treatment effect measures. Standardization of procedures for student selection must be viewed from these two points of view.

The procedures for generating the Norm-Referenced Model no-treatment expectation are:

1. Select students for the Title I project on some measure other than their pretest score.
2. Pretest Title I students with a nationally-normed test at a time corresponding to empirical norm dates:
3. Compute the mean pretest standard score for Title I students (excluding students who do not have posttest scores).
4. Convert the mean pretest standard score to its percentile equivalent.
5. Define the percentile equivalent of the mean pretest standard score as the no-treatment expectation for the Title I group.

Separating the pretest from the selection measure is intended to control regression effects. If this control is inadequate, then the equipercentile assumption applied in step 5 appears to be false.

Initial work addressed the question, 'Do independent administrations of the selection test and the pretest control for regression effects in the Norm-Referenced Model when the mean pretest of the selected group is used to define the no-treatment expectation?' Analysis implied that implementing the Norm-Referenced Model may result in a biased treatment effect. Independent selection and pretest measures reduce but do not necessarily control for regression. Bias is indicated when the correlation between the selection measure and the pretest is different from the correlation between the selection measure and the posttest--$r_{x_0x_1} \neq r_{x_0x_2}$. Where:

X_0 denotes selection test scores obtained at occasion 0,

X_1 denotes pretest scores obtained at occasion 1, and

X_2 denotes posttest scores obtained at occasion 2.

In most cases, the inequality will be $r_{x_0x_1} > r_{x_0x_2}$. The selection test will correlate higher with the pretest than it will with the posttest. When $r_{x_0x_1} > r_{x_0x_2}$, regression bias will result even though the selection test scores are obtained independently of the pretest scores. The direction of bias will result in an overestimate of the Title I treatment effect.

After demonstrating that the Norm-Referenced Model would likely produce biased treatment effect measures, four possible actions were considered--they were:

1. Recommend dropping the Norm-Referenced Model from the System.
2. Recommend treating results of Norm-Referenced Model evaluations

separately from those produced by the Control Group Model and the Special Regression Model and qualify conclusions reached in Norm-Referenced evaluations.

3. Encourage SEA's and LEA's to choose another model.

4. Develop a sounder method for estimating the Norm-Referenced Model no-treatment expectation.

The purpose of present analysis is to provide the theoretical basis for improving the no-treatment expectation estimation procedures for the Norm-Referenced Model. The method used to derive the alternative definition of the no-treatment expectation is based on a structural equation model. The model developed is an extension of a model Kenny (1975) developed for analyzing nonequivalent control group designs.

Structural equation models—in this case a recursive path analysis model—can be used to derive the appropriate no-treatment expectation for the nonequivalent control group design where student selection has been based on a separate measure of the same achievement construct measured by the pretest and posttest measures. The model, which may also be called a causal model, includes a clear statement of the distributional assumptions and causal ordering of the variables assumed to underlie the measures obtained in the evaluation. The causal model used to derive an alternative no-treatment expectation for the Norm-Referenced Model does not require the assumption of equipercentile growth or the assumption that regression is controlled by the dual pretest design.

In order to intuitively grasp the mathematical argument which follows, consider Kenny's advice that:

"the researcher carefully study the process of selection into treatments.

The critical questions are: (a) With what and by how much does the treatment correlate with the causes of the dependent variable?, and (b) When does selection take place? The answer to these questions determines the mode of analysis." (page 359)

In other words, if the process of selecting Title I students from a known population (or populations) can be quantitatively modeled, one can rationally generate a no-treatment expectation.

Random assignment of students from a single population is a selection process with well known properties. It supplies the theoretical basis for the Control Group Model no-treatment expectation with no adjustment and the analysis of covariance adjustment. Random (or complete) assignment of students to Title I from one population and random (or complete) assignment of students to a Control Group from a second population provides the theoretical basis for the principal axis adjustment in the Control Group Model (Kenny, 1975).

Likewise, selecting Title I students from a single population on the basis of the best available covariate (e.g., the pretest score or a composite measure including the pretest score) provides theoretical justification for establishing the Special Regression Model no-treatment expectation. Theoretically, the adjustment applied in the Special Regression Model is identical to the analysis of covariance adjustment for the Control Group Model. However, the procedures for estimating the adjustment factors are model-specific.

Before presenting the adjustments derived by Kenny, treatment effect correlations need to be briefly explained because the structural equation model includes the treatment variable as a dummy variable.

Treatment Effect Correlation

To understand the statistical argument in this paper it will be necessary to conceptualize treatment effect measures in correlational terms as well as mean difference terms. The null hypothesis for an evaluation model that contrasts a treated group with an untreated group can be expressed as the expected point-biserial correlation between the treatment (e.g., coded as a dummy variable where treatment group members are scored as 1 and control group members are coded as 0) and the posttest score. The magnitude of the correlation is a function of (a) the mean difference between the treated group and the untreated group on the continuously distributed pretest or posttest scores, (b) the proportion of subjects in the treated or untreated group, and (c) the standard deviation of the combined groups on the continuous variable.

Two alternative formulas for the point-biserial correlation are:

$$r_{ptbis} = \frac{\bar{X}_t - \bar{X}_c}{\sigma_g} \cdot \sqrt{pq} \tag{1}$$

--OR--

$$r_{ptbis} = \frac{\bar{X}_t - \bar{X}_g}{\sigma_g} \cdot \sqrt{\frac{p}{q}} \tag{2}$$

Where:

t represents the treatment group,
c represents the control group,
g represents the combined treatment and control group,
p represents the proportion of subjects in the g who are in treatment group,
q represents the proportion of subjects in the g who are in the control group,
σ_g represents the standard deviation of the g on the continuous variable (i.e., pretest or posttest),
$\bar{X}_t$ represents the mean of the treatment group (t) on the continuous variable,
$\bar{X}_c$ represents the mean of the control group (c) on the continuous variable,
$\bar{X}_g$ represents the mean of the combined group (g) on the continuous variable.

The impact of a treatment can be assessed by comparing two treatment effect correlations--the first is based on treatment status correlated with pretest status (i.e., the pretreatment effect correlation) and the second is based on treatment status correlated with posttest status (i.e., the posttreatment effect correlation). Assuming that σ_g, p and q remain constant, a treatment effect may be indicated by a more positive (or less negative) posttreatment effect correlation compared to the pretreatment effect correlation.

When a randomized experiment has been implemented successfully and the means of the treated group and the untreated group on the pretest are equal, the pretreatment effect correlation is zero. In expectancy terms, the expected value of the pretreatment effect correlation under random assignment is zero.

When randomization has failed to produce equal pretest means due to chance, or where selection is based on factors correlated with the pretest, the pretreatment effect correlation will be nonzero. If Title I students are selected on the basis of low educational achievement and higher achievers make up the comparison group, the pretreatment effect correlation (r_{X_1T}) will be negative.

A no-treatment expectation, or a null hypothesis, can be expressed as an expected value of the posttreatment effect correlation (r_{X_2T}). An expected null posttreatment effect correlation can also be converted directly to an expected mean value on a posttest measure for the treatment group, assuming the treatment had no effect. In other words, by solving for $\bar{X}_{t2}$ in the formula for a point-biserial correlation, the no-treatment expectation can be expressed as a mean posttest score in an appropriate metric. The regression of X_2 on T is used to predict the value of X_2 for those in the treatment group.

Solving for $\bar{X}_{t2}$ is illustrated in the following example.

Using formula 3,

$$r_{x_i T} = \frac{\bar{X}_{ti} - \bar{X}_{gi}}{\sigma_g} \sqrt{\frac{p}{q}} \tag{3}$$

Where i represents the occasion of testing (i.e., pretest or posttest),

where $\bar{X}_{ti}$ represents the mean score of the treatment group on the continuous variable (e.g., mean posttest score of Title I group),

where $\bar{X}_{gi}$ represents the mean score of the combined groups on the continuous variable (e.g., mean score of local norm group taken at the time of the posttest).

Solving for $\bar{X}_{2t}$:

$$r_{x_2 T} = \frac{\bar{X}_{t2} - \bar{X}_{g2}}{\sigma_g} \cdot \sqrt{\frac{p}{q}}$$

$$\frac{r_{x_2 T}\sigma_g}{\sqrt{p/q}} = \bar{X}_{t2} - \bar{X}_{g2}$$

$$\frac{r_{x_2 T}\sigma_g}{\sqrt{p/q}} + \bar{X}_{g2} = \bar{X}_{t2} \tag{4}$$

Given a hypothesized value of $r_{x_2 T}$, an expected value of $\bar{X}_{t2}$ may be computed under the null hypothesis.

<u>Example</u>. Assume a hypothesized no-treatment effect correlation with a value of:

$$r_{x_2 T} = -.40.$$

This value means that the treatment group is expected to score lower on the posttest than the combined group (e.g., a local norm group).

Second, assume a value of 21.06 for the standard deviation of the combined group on the continuous achievement measure:

$$\sigma_g = 21.06.$$

Third, assume 30 percent of the students in the combined group are in the treatment group and 70 percent of the students are in the nontreatment group. Thus,

$$p = .30 \text{ and } q = .70.$$

Fourth, assume that the mean score of the combined group at the time of the posttest is 50 NCE's (a standard score used in Title I Evaluation and Reporting System):

$$\bar{X}_{g2} = 50.$$

Then, in solving for $\hat{\bar{X}}_{t2}$, the predicted posttest mean for the treatment group where treatment has had no effect, we have:

$$\hat{\bar{X}}_{t2} = \frac{.40\ (21.06)}{.30/_{.70}} + 50 = \frac{-8.42}{.655} + 50 = 37.1.$$

The value 37.1 is the no-treatment expectation expressed as the predicted mean of the treatment group on the posttest.

This example illustrates how a hypothesized no-treatment expectation expressed as a point-biserial can be converted to a no-treatment expectation expressed as a mean. The Title I no-treatment expectations, then, can be expressed (under specific conditions) either as a mean or a posttreatment effect correlation. Therefore, the proof included in this paper applies directly to the Title I Evaluation and Reporting System convention for expressing the no-treatment expectation.

General Formulas for Expressing Null Hypotheses

As indicated above, the appropriate no-treatment expectation for the randomized experiment is expressed as follows:

$$H_0:\ r_{x_1T} = r_{x_2T} = 0;$$

or, in terms of mean scores:

$$H_0:\ \mu_{t2} = \mu_{c2}.$$

The problem arises when $r_{x_1T} \neq 0$, or where $\bar{X}_{t1} \neq \bar{X}_{c1}$, which may be the result of chance or a nonrandom selection process. A statistical adjustment may be employed to generate the appropriate null hypothesis in this instance.

Kenny (1975) has expressed the null hypothesis associated with each of four common alternative estimation methods. They are presented in Table 1.

TABLE 1

Null Hypotheses Corresponding to Four Statistical Estimation Methods

ESTIMATION METHOD	NULL HYPOTHESIS
ANCOVA	$r_{x_2T} = r_{x_1T}r_{x_1x_2}$
ANCOVA with Reliability Correction	$r_{x_2T} = r_{x_1T}r_{x_1x_2}/r_{x_1x_1}$
Raw Change Score Analysis	$r_{x_2T} = r_{x_1T}\sigma_{x_1}/\sigma_{x_2}$
Standardized Change Score Analysis	$r_{x_2T} = r_{x_1T}$

Using a general linear model and a path analytic logic, Kenny specified three alternative selection processes and derived the appropriate null hypothesis from each of them. These selection processes and their associated null hypotheses are presented in Table 2.

TABLE 2

Null Hypotheses Corresponding to Four Selection Processes

SELECTION PROCESS	NULL HYPOTHESIS
Based on the Pretest	$r_{x_2T} = r_{x_1T} r_{x_1x_2}$
Based on Pretest True Score	$r_{x_2T} = r_{x_1T} r_{x_1x_2} / r_{x_1x_1}$
Based only on Group Membership	$r_{x_2T} = r_{x_1T}$
Between the Pretest and Posttest	$r_{x_2T} = r_{x_1T}$

Kenny's model can be used to derive the no-treatment expectations for the Control Group Model with no adjustment, the ANCOVA adjustment and the principal axis adjustment. It can also be used to generate the no-treatment expectation for the Special Regression Model. Since the Kenny model includes only one pretest score, however, it cannot be used to derive the no-treatment expectation for the Norm-Referenced Model. In this paper, a second pretest variable and two unmeasured variables are added so that a no-treatment expectation for the Norm-Referenced Model can be derived. Thus, the basis for the Norm-Referenced Model no-treatment expectation is made explicit in its definitions and underlying assumptions. Moreover, the newly-defined model also provides a basis for comparing all three Title I evaluation models (including the three adjustment alternatives for the Control Group Model).

Derivation of a No-Treatment Expectation for the Norm-Referenced Evaluation Model

While the Norm-Referenced Evaluation Model does not specify standardized procedures for selecting students into the Title I program, it is a common practice to use some measure of individual student achievement in the selection process. This practice is supported by Title I regulations and recommended by RMC. It would be informative, then, to develop a formal psychometric and causal model that would result in a no-treatment expectation where subjects are selected into Title I on the basis of a measure taken prior to the pretest. The model includes four measured variables--they are:

X_0 = an achievement measure used for selection,

X_1 = a pretest measure of achievement,

X_2 = a posttest measure of achievement,

T = the treatment variable.

The causal function of X is based on three unmeasured hypothetical constructs:

G = group membership (e.g., sex, race and classroom),

Z = individual differences within groups (e.g., ability, achievement),

E = errors of measurement or unstable causes of Z.

Assuming all variables have been standardized to reduce algebraic complexity, we can express X_1 and X_2 in terms of their causes, G, Z and E:

$$X_{0i} = a_0G_i + b_0Z_{0i} + e_0E_{0i}, \tag{5}$$

$$X_{1i} = a_1G_i + b_1Z_{1i} + e_1E_{1i}, \tag{6}$$

$$X_{2i} = a_2G_i + b_2Z_{2i} + e_2E_{2i}, \tag{7}$$

where the subscripts 0, 1 and 2 refer to occasions, and subscript i to the subject.

Note that, since we are deriving the no-treatment expectation, no effects on X_{21} are assumed to result from T, the treatment variable.

The following assumptions are made for this series of linear equations:

1. Group membership (G) is assumed to remain constant across time; its autocorrelation equals 1.00.
2. Relative position within groups (Z) <u>may</u> vary across time; its autocorrelation may be less than 1.00.
3. Errors of measurement (E) are perfectly unstable, or random; its autocorrelation is 0.
4. All unmeasured variables (G, Z and E) are assumed to be uncorrelated with each other.
5. Group membership (G) and relative position within groups (Z) do not interact.

The first two assumptions mean that an individual's status as a member of some group remains constant over time, while status on individual traits varies over time. According to the model, Z is defined as relative position <u>within</u> groups (e.g., standard scores based on differentiated norms would represent a Z), and errors of measurement are defined as uncorrelated with other errors of measurement, as well as with G and Z. It follows then that there is no confounding of effects of the unmeasured variables on X. Finally, the assumption that G and Z do not interact is analogous to the assumption of homogeneity of regression in analysis of covariance. It means that the regression of X on Z is the same for all groups considered.

If the treatment is correlated with the selection test, then it must be confounded with the causes of the scores on the selection test. In other words, whether an individual is in the treatment group or the control group is a function of his group membership, his relative status within the group, errors of measurement or some combination of the three variables. Given this, the causal function for the treatment variable is written as:

$$T_i = qG_i + mZ_{0i} + sE_{0i} + fU; \tag{8}$$

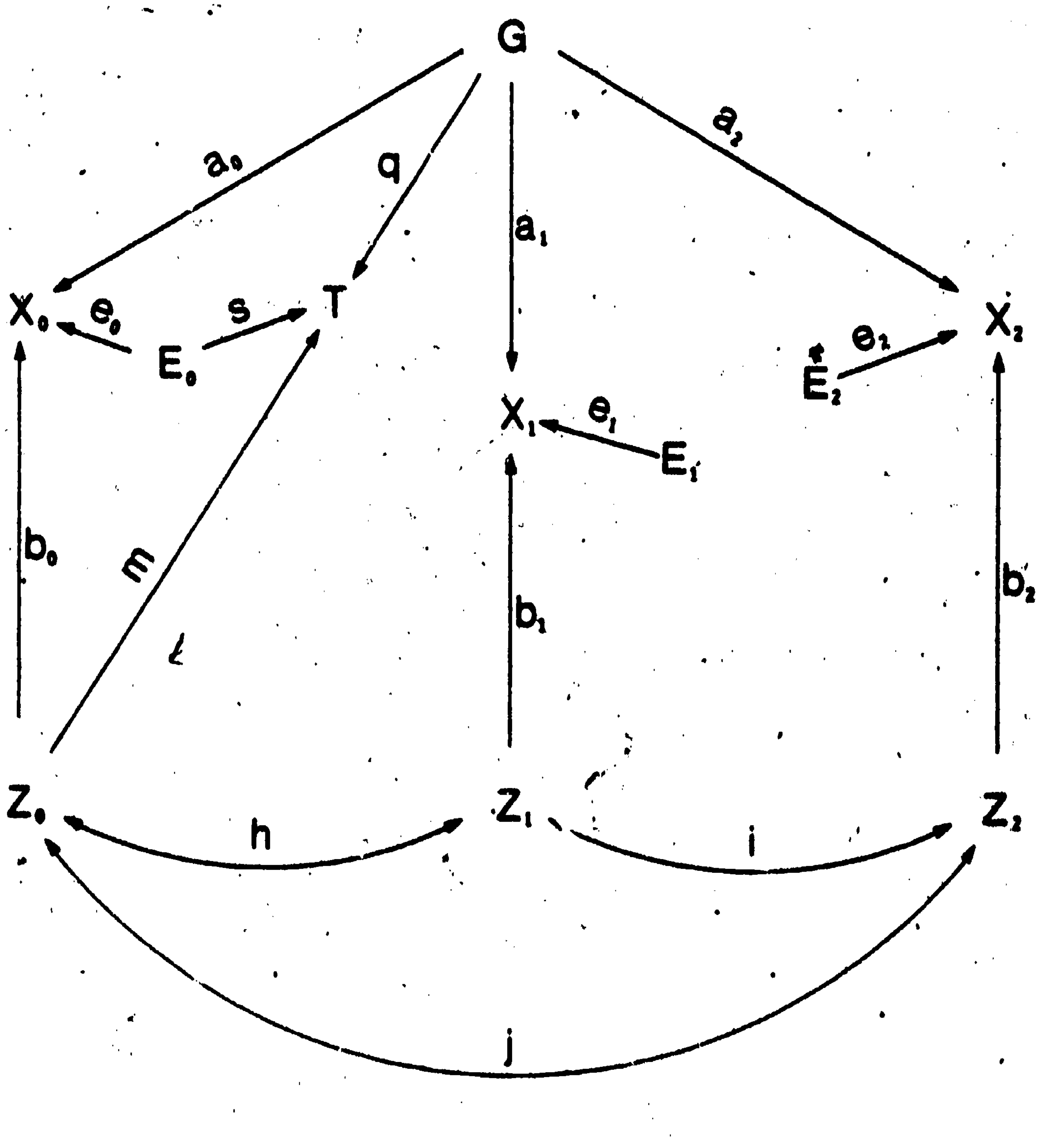
where U is a residual term that is uncorrelated with all unmeasured variables. The smaller the residual term, the more reasonable the model that shows treatment and control group membership to be based on the same variables that determine pretest status. A large residual term would indicate that assignment to treatment is based upon variables that are uncorrelated with G, Z and E. Since these variables are unmeasured, a direct assessment of the residual term is not made.

The next step in developing the no-treatment expectation is to analyze the correlations between the measured variables X_0, X_1, X_2 and T. To do this, we first express the correlations (which are observable) in terms of the standardized partial regression coefficients (which are unobservable) specified in equations 5, 6, 7 and 8. The principles of path analysis aid in this process. Accordingly, we have provided a path diagram (Figure 1) for the model specified in equations 5 through 8 and the assumptions which support those equations.

16

FIGURE 1

Path Diagram Depicting the Application of Independently Administered Tests for Selection (X_0) and Pretesting (X_1)



Since G, 2 and E are latent variables, the regression coefficients linking those variables with the measured variables X_0, X_1, X_2 and T are also unobservable. However, by making certain assumptions and taking advantage of algebraic identities, we can use these coefficients to show theoretically-expected relationships among the measured variables.

The correlations between the observed variables (X_0, X_1, X_2 and T) are expressed in terms of the regression coefficients (path coefficients) in Figure 1. These correlations are:

$$r_{x_0T} = qa_0 + mb_0 + se_0 \tag{9}$$

$$r_{x_1T} = qa_1 + mhb_1 \tag{10}$$

$$r_{x_2T} = qa_2 + mjb_2 \tag{11}$$

$$r_{x_0x_1} = a_0a_1 + b_0b_1h \tag{12}$$

$$r_{x_0x_2} = a_0a_2 + b_0b_2j \tag{13}$$

$$r_{x_1x_2} = a_1a_2 + b_1b_2i \tag{14}$$

In each of these equations, the correlation between two variables is taken as the sum of the products of certain "path coefficients." For those who are more familiar with multiple regression techniques, each of these path coefficients can be shown in more conventional terms as standardized partial regression coefficients.

Having expressed the correlations between the measured variables in terms of path coefficients, we now proceed to further specify the model in terms which will eventually allow us to express expected relationships between the observed correlations. What we are moving toward is a way to express r_{x_2T} as a function of r_{x_1T}.

If selection is based upon X_0, we have:

$$T_i = kX_{0i} + fU_i, \qquad (15)$$

which expresses membership in the treatment group as a function of the selection test score (X_0) and a residual (U).

Next, we can substitute equation 5 for X_{0i} in equation 15, which gives:

$$T_i = k\,(a_0G_i + b_0Z_{0i} + e_0E_{0i}) + fU_i.$$

Since T also equals $qG_i + mZ_{0i} + sE_{0i} + fU_i$ from equation 8, we have:

$$a_0k = q, \quad b_0k = m, \quad e_0k = s,$$

—OR—

$$q/a_0 = m/b_0 = s/e_0 = k.$$

Knowing that $a^2_0 + b^2_0 + e^2_0 = 1$ because X_i is standardized, it is possible to show r_{x_0T} is equal to k:

$$r_{x_0T} = qa_0 + mb_0 + se_0 \quad \text{(equation 9)}$$

$$r_{x_0T} = a_0ka_0 + b_0kb_0 + e_0ke_0$$

$$r_{x_0T} = a^2_0k + b_0{}^2k + e_0{}^2k$$

$$r_{x_0T} = k$$

Next, we show that $r_{x_1T} = kr_{x_0x_1}$ since $q = a_0k$ and $m = b_0k$:

$$r_{x_1T} = a_0ka_1 + b_0khb_1 \quad \text{(equation 10)}$$

$$r_{x_1T} = k\,(a_0a_1 + b_0b_1h)$$

$$r_{x_1T} = kr_{x_0x_1}$$

Now that we have $r_{x_0T} = k$ and $r_{x_1T} = kr_{x_0x_1}$, we can substitute for k in $kr_{x_0x_1}$ and find:

$$r_{x_1T} = r_{x_0T} \cdot r_{x_0x_1}.$$

This expresses the correlation between the pretest and the treatment in terms of the correlation between the selection test and treatment and the correlation between the selection test and the pretest. Since $r_{x_0x_1}$ will nearly always be less than 1.00, we can expect r_{x_1T} to be less than r_{x_0T}. The equation $r_{x_1T} = r_{x_0T} \cdot r_{x_0x_1}$ is identical to the null hypothesis for ANCOVA if X_1 were the posttest rather than the pretest. The solution to this point is identical to the solution offered by Kenny (1975), page 350. Since the logic of Model A requires that we solve for r_{x_2T} in terms of r_{x_1T}, additional steps are needed to find these results.

We begin by showing that:

$$r_{x_2T} = a_0ka_2 + b_0kjb_2,$$

which follows directly from $r_{x_2T} = qa_2 + mjb_2$, and $q = a_0k$, and $m = b_0k$.

Factoring k out of $a_0ka_2 + b_0kjb_2$ gives:

$$r_{x_2T} = k\,(a_0a_2 + b_0jb_2).$$

Knowing that:

$$r_{x_0x_2} = a_0a_2 + b_0jb_2 \quad \text{(equation 13)}$$

leads to:

$$r_{x_2T} = kr_{x_0x_2}.$$

Substituting for k gives:

$$r_{x_2T} = r_{x_0T} \cdot r_{x_0x_2};$$

but, since $r_{x_0T} = r_{x_1T} / r_{x_0x_1}$, because $r_{x_1T} = r_{x_0T} \cdot r_{x_0x_1}$, we have:

$$r_{x_2T} = r_{x_1T}\,(r_{x_0x_2} / r_{x_0x_1}). \tag{16}$$

This final expression gives the expected value of r'_{x_2T}, where:

(1) A selection test has been used to assign students to treatment;

(2) An independently administered test has been used as a pretest, which is then used to generate a baseline or no-treatment expectation.

The ratio $r_{x_0x_2}/r_{x_0x_1}$ is the appropriate statistical adjustment for r_{x_1T}, assuming that regression is homogeneous and linear.

An Example Comparing Alternative No-Treatment Estimation Methods for the Norm-Referenced Model

A Title I program testing schedule may include a selection test in the spring, a pretest in the fall, and a posttest in the following spring. Table 3 shows the uses that might be made of these test data in a Title I evaluation.

TABLE 3

Use of Test Data

TIME OF TESTING	PLAN 1	PLAN 2	PLAN 3
Spring '78	selection	not used	selection and pretest
Fall '78	pretest	selection and pretest	not used
Spring '79	posttest	posttest	posttest

Each column represents a plan for test use. The pretest is used to generate the no-treatment expectation in each of these cases. A separate selection test is applied in Plan 1.

Three statistical estimation factors that may be applied to these plans are:

$\sigma_{x_2}/\sigma_{x_1}$ (Method A)

(assume variances are equal so $\sigma_{x_2}/\sigma_{x_1} = 1.00$);

$r_{x_0x_2}/r_{x_0x_1}$; (Method B)

$r_{x_1x_2}$. (Method C)

Method A corresponds to the standardized gain or principal axis estimation factor. When NCE (Normal Curve Equivalent) scores are used (and $\sigma_{x_2} = \sigma_{x_1}$), Method A corresponds to the equipercentile growth estimation technique which RMC recommends for the Norm-Referenced Model.

Method B is the statistical estimation factor derived in this paper as the method of choice when selection has been based on an achievement test other than that used as the pretest. Method C, on the other hand, is the estimation factor defined by ANCOVA. The ANCOVA estimation factor is applied here, as it has been recommended as an alternative to using the equipercentile assumption in the Norm-Referenced Model.

Since Plan 1 is the Norm-Referenced design recommended by RMC, the most significant comparison is between Method A and Method B applied to Plan 1.

Suppose we had the following correlations as population parameters:

$r_{x_0x_1} = .80,$

$r_{x_0x_2} = .70,$

$r_{x_1x_2} = .80;$

where X_0 is Spring '78 test score,

X_1 is Fall '78 test score,

X_2 is Spring '79 test score.

Next, suppose we had selected students on X_0 so that the mean score of the selected group was one standard deviation below the population mean. The mean of the selected group in terms of Z scores and NCE's would be:

$$Z_{t_0} = -1.00 \quad \text{(an NCE of 28.94).}$$

If we use Z_{t_0} and $r_{x_0x_1}$ (again assuming $\sigma_{x_0}=\sigma_{x_1}$) to provide the estimate $\hat{Z}_{t_1}$, we find:

$$\hat{Z}_{t_1} = -.80 \quad \text{(an NCE of 33.15)}$$

Given the values of $r_{x_0x_2}$, $r_{x_0x_1}$, $r_{x_1x_2}$, Z_{t_0} and $\hat{Z}_{t_1}$, we can apply Methods A, B and C to Plan 1. The results are presented in Table 4.

TABLE 4

No-Treatment Expectations Expressed in NCE's: Plan 1

METHOD	NO-TREATMENT EXPECTATION
Method A	33.15
	Difference = 2.11
Method B	35.26
	Difference = 1.26
Method C	36.52

The no-treatment expectations in Table 4 were found by applying the following two formulas:

$$\hat{Z}_{tx_2} = \text{Estimation Factor} \cdot \hat{Z}_{tx_1},$$

where $\hat{Z}_{tx_1}$ is the expected mean Z score of the treatment group on the pretest,

Then the value of $\hat{Z}_{tx_2}$ was transformed to an NCE by:

$$\text{Mean NCE of Treated Group, assuming no effect} = 50.00 + 21.06\,(\hat{Z}_{tx_2}).$$

The results show that the equipercentile growth method recommended by RMC <u>underestimates</u> the no-treatment expectation by 2.11 NCE's. On the other hand, ANCOVA with the pretest as the covariate <u>overestimates</u> the no-treatment expectation by 1.26 because prior selection on X_0 was ignored in choosing the value of $r_{x_1x_2}$ as the estimation factor.

The amount of regression from the selection test to the pretest is from 28.94 to 33.15, or 4.21 NCE's. The amount of regression from the pretest to the posttest is from 33.15 to 35.26, or 2.11 NCE's. The RMC method accounted for only two-thirds of the regression, whereas alternative Method B accounted for all the regression.

One could argue that Plans 2 and 3 also be considered as possible in the Norm-Referenced Model. Table 5 shows the result of applying Method A (this method is not recommended by RMC unless a prior selection test has been used) and Method C.

TABLE 5

No-Treatment Expectations in NCE's: Plans 2 and 3

METHOD	NO-TREATMENT EXPECTATION	
	PLAN 2	PLAN 3
Method A	28.94	28.94
Method C	33.15	35.26

Since selection was based on the pretest in Plans 2 and 3, the mean pretest score of the Title I treatment group in Z score terms is -1.00. However, the occasion for the pretest in Plan 2 is the Fall of 1978, whereas the occasion for the pretest in Plan 3 is the Spring of 1978.

Method A is not recommended for use in Plans 2 or 3. In this example, Method A compared to Method C illustrates the amount of regression operating across two time intervals. Method C applied to Plan 3 gives the same no-treatment expectation as Method B does in Plan 1, as expected.

Summary and Discussion

Using a measure other than the pretest score for selecting students, as in the Norm-Referenced Model, reduces but does not control regression effects. If the selection measure is more highly correlated with the pretest measure than it is with the posttest measure, additional regression occurs between the pretest and the posttest. When an extreme group has been selected for treatment, the effects of regression will bias treatment effect measures unless appropriate methods for estimating the no-treatment expectation are applied.

As it is applied in the Norm-Referenced Model, the equipercentile growth assumption does not provide a sound unbiased estimation method. This follows directly from proof that the dual pretest procedure, with one test used for selection and the second used for estimation under the equipercentile assumption, does not control for regression.

Having established a problem with the dual pretest design, four general solutions are suggested. Possible solutions to the problem include:

1. Recommend dropping the Norm-Referenced Model from the USOE Title I Evaluation and Reporting System.
2. Recommend that results of the Norm-Referenced Model not be aggregated with those of other models and that conclusions based on the Norm-Referenced Model be subject to special qualifications.
3. Encourage SEA's and LEA's to choose another model.
4. Develop a sound method for estimating the Norm-Referenced Model no-treatment expectation.

The general solution presented in this paper is to develop a sound method for estimating the Norm-Referenced Model no-treatment expectation. Recent work on quasi-experimentation stresses that failure to randomly assign subjects to treatment and control groups does not necessarily leave one unable to estimate treatment effects free from bias. Where the selection process is known to the extent that it can be mathematically modeled, it is still possible, in theory, to provide unbiased estimates of treatment effects (Cain, 1977; Rubin, 1977; Kenny, 1975; Bryk and Weisberg, 1977). Cronbach, Rogosa, Ploden and Price (1976) point out that it is quite difficult to effectively control the parameters which do affect bias.

The specific solution derived is theoretical. It applies where student selection is based on a measure of the same construct as a separately administered pretest and posttest. While the analytical model used to derive the no-treatment expectation involved a two-factor model underlying the achievement measures, the same solution follows from a single-factor model where the group membership factor is ignored.

The important empirical facts that the model takes into account include the extremity of the treatment group mean from the population mean and the extent of temporal erosion.

The first empirical consideration, the extremity of the treatment group, can be measured by the pretreatment effect correlation, r_{x_1T}. If the pretreatment effect correlation is zero, the treatment group is not extreme and regression effects would not be expected. In other words, if the treatment group's pretest mean is the same as the population's pretest mean, then no regression effects would be anticipated.

The second empirical consideration, the amount of temporal erosion, is indicated by the ratio $r_{x_0x_2}/r_{x_0x_1}$. The greater the amount of temporal erosion, the lower the value of this ratio and the greater the expected regression. When there is no temporal erosion, the value of $r_{x_0x_2}/r_{x_0x_1}$ will be 1.00 and no regression would be expected, regardless of the extremity of the treatment group. The effect of the ratio is to apply the appropriate weight to r_{x_1T} in estimating the expected value r_{x_2T}.

These observations apply only to the dual pretest design implemented as in the RMC version of the Norm-Referenced Model.

Measuring Treatment. The analytical model used to derive the alternative no-treatment expectation presented in this paper assumes clearly defined treatment and comparison groups. If a standard nonequivalent comparison group design is used, then two groups can be identified from a single population and the measure of T is straightforward. The Norm-Referenced Model, however, uses norms tables

as a source of population parameter estimates instead of a concrete comparison group. This is probably seen as a practical virtue of the Norm-Referenced Model. The practical advantage takes a toll in the assumptions required in order to make the Norm-Referenced Model workable.

The measure of the pretreatment effect correlation from which the estimated no-treatment effect correlation is derived simply contrasts the treatment group's pretest mean to the mean of the population from which it was drawn. Here lies the rub. The choice of norms determines the extremity of the treatment group. If the treatment group can be considered a sample from the population represented in national norms (or local norms), then those national (or local) norms supply the population parameter estimates needed. These parameters include mean scores and standard deviations for the two points in time corresponding to the pretest and posttest dates and the slopes of the regression lines predicting pretest scores from selection test scores and posttest scores from pretest scores.

Thus, the problem of choosing appropriate norms (e.g., local versus national norms) is a problem of determining appropriate procedures for estimating the parameters needed. It is a theoretical or logical problem as well as an empirical problem. Using an achievement measure to select treatment students virtually guarantees that an extreme group (one whose mean differs from the population mean will result no matter how complex the definition of the population is.

Use of Ex Post Facto ANCOVA. The dual pretest design precludes applying the equipercentile growth assumption as a quantitatively simple procedure for estimation. Consequently, we must question the need for the separate selection test. Might not

the pretest be used for selection and estimation of the no-treatment expectation, using $r_{x_1x_2}$ to make the prediction of r_{x_2T}? The answer, assuming students are drawn from a single known population, is yes, but this answer holds only if prior selection (i.e., prescreening) has not taken place. The analysis in this paper shows ANCOVA to underestimate treatment effects when prescreening is ignored. Multi-stage selection processes must be taken into account in estimating.

Using norms could represent an extremely practical approach to estimating the treatment effects of educational programs. The procedures for sound use of norms must be carefully thought out and studied. Development of the model in this paper is a further st p in improving theoretical understanding of the Title I Evaluation and Reporting System. The assumptions behind the model are parsimonious. hey may not be met in practice. If they are not met, then we would expect additional erro. to enter into the System. However, departures from normality, linearity and homogeneity of regression will not likely make it procedurally or theoretically easier to estimate the no-treatment expectation in norm-referenced quasi-experimental design. In all probability, such departures will make estimation procedures more complex.

References

Bryk, A.S., and Weisberg, H.I. Use of the nonequivalent control group design when subjects are growing. Psychological Bulletin, 1977, 84, pp. 950-962.

Cain, G.C. Regression and selection models to improve nonexperimental comparisons. In M. Guttentag (Ed.), Evaluation Studies Review Annual. Beverly Hills, 1977: Sage Publications.

Cronbach, L.J.; Rogosa, D.R.; Ploden, R.E.; and Price, G.G. Analysis of covariance in nonrandomized experiments: Parameters affecting bias. Stanford Evaluation Consortium, Occasional Paper, 1976.

Kenny, D.A. A quasi-experimental approach to assessing treatment effects in the nonequivalent control group design. Psychological Bulletin, 1975, 82, pp. 345-362.

Rubin, D.B. Assignment to treatment group on the basis of a covariate. Journal of Educational Statistics, 1977, 2, pp. 1-26.

Tallmadge, G.K., and Wood, C.T. User's guide, ESEA Title I evaluation and reporting system (revised). Prepared for USDHEW/USOE Office of Planning, Budgeting and Evaluation. Mountain View, CA.: RMC Research Corporation, December 1976.